Information Theory, Fall 2015	Iftach Haitner
Problem set 2	
December 8, 2015	Due: Nov 24

- Please submit the handout in class, or email the grader (omer.rotem1 at gmail.com).
- Write clearly and shortly using sub-claims if needed. The emphasize in most questions is on the proofs (no much point is writing a “solution" w/o proving its correctness)
- For Latex users, a solution example can be found in the course web site.
- It is allowed to work in (small) groups, but please write the id list of your partners in the solution file, and each student should write his solution by *himself* (joint effort is only allowed in the “thinking phase")

1. Prove the chain rule from mutual information (see, Lecture 2, Slide 21)
2. For random variables X and Y, what is larger $I(X;Y|Z)$ or $I(X;Y)$?
3. Let $X \to Y \to Z$ be a Markov chain. Is it always the case that $I(Y;Z) \geq I(X;Z)$?
4. Formally define and prove the infinite case of craft inequality (Thm. 2, Lecture 4).
5. Let $X \sim (p_1, \ldots, p_m)$ such that each p_i is a power of 2 (i.e., 2^{-k} for some $k \in \mathbb{Z}$).
   (a) (This part did not appear in the version I asked you to solve, but it is need for solving Q6).
   Prove that Huffman's code assign a word $x \in \mathrm{Supp}(X)$ of probability 2^{-i}, a codeword of length i.
   (b) Prove that the average code length obtained by Huffman's code for X is (exactly) $H(X)$.
6. Use the above question and the optimality of Huffman's code to deduce that the average code length obtained by Huffman's code on any random variable X is at most $H(X) + 1$. (Do that without using the upper bound on the optimal code length proved in class).
   Can you prove it without the assumption that the binary expansion of each p_i is finite?
7. Prove Proposition 5 in Lecture 4.